

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-334070

(43)Date of publication of application : 18.12.1998

(51)Int.Cl.

G06F 15/18

(21)Application number : 09-144331

(71)Applicant : NIPPON TELEGR & TELEPH CORP <NTT>

(22)Date of filing : 02.06.1997

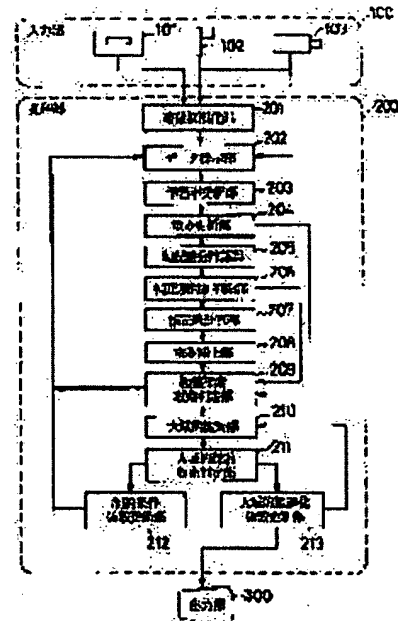
(72)Inventor : OCHIAI YOSHIHIRO
SUZUKI SATOSHI

(54) METHOD AND DEVICE FOR GLOBAL ACCELERATION LEARNING FOR NEURAL CIRCUIT NETWORK MODEL

(57)Abstract:

PROBLEM TO BE SOLVED: To shorten the search time by repeatedly executing a global search when a solution having an evaluation function value smaller than the evaluation function value of local solution can not be searched.

SOLUTION: A correction amount calculation part 207 calculates the correction amount of weight from the difference of weight or gradient transferred according to the command of correction judge part 206, a weight correction part 208 corrects the value of weight, and an extremal value search convergence judging part 209 judges whether the stop preference of extremal value search is satisfied or not. A global search part 210 executes minimization to an evaluation function and searches a point where the evaluation function value gets smaller. A global search convergence judging part 211 judges whether the point found as a result of global search satisfies the stop recurrence of macro optimization on not. When the point of smaller evaluation function value is found out, a limit condition coefficient change part 212 changes the coefficient of limit condition. When the point of smaller evaluation function value is not found out, a global optimization coefficient change part 213 updates the coefficient of global optimization and when each processing is not finished even after repetition, processing is finished.



LEGAL STATUS

[Date of request for examination] 26.12.2000

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the
examiner's decision of rejection or application
converted registration]

[Date of final disposal for application]

[Patent number] 3597341

[Date of registration] 17.09.2004

[Number of appeal against examiner's decision of
rejection][Date of requesting appeal against examiner's decision
of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-334070

(43) 公開日 平成10年(1998)12月18日

(51) Int.Cl.⁶
G 0 6 F 15/18

識別記号
5 2 0

F I
G 0 6 F 15/18

5 2 0 Q

審査請求 未請求 請求項の数 6 O L (全 18 頁)

(21) 出願番号 特願平9-144331

(22) 出願日 平成9年(1997)6月2日

(71) 出願人 000004226

日本電信電話株式会社
東京都新宿区西新宿三丁目19番2号

(72) 発明者 落合 慶広

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72) 発明者 鈴木 智

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

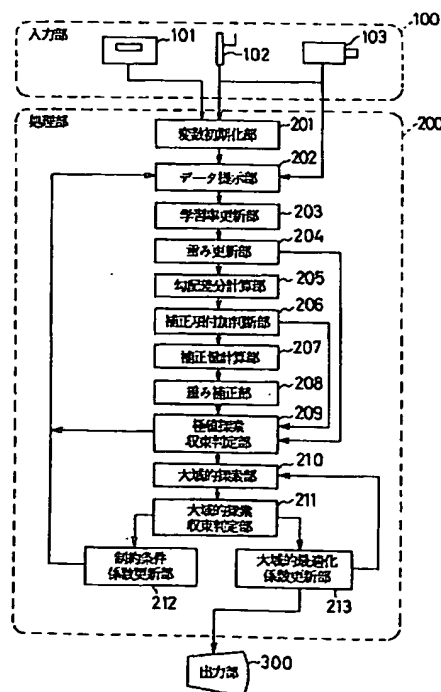
(74) 代理人 弁理士 三好 秀和 (外1名)

(54) 【発明の名称】 神経回路網モデルの大域的加速化学習法およびその装置

(57) 【要約】

【課題】 評価関数曲面上に局所的極小値が存在しても大域的な極小値を探索することができると共に学習の収束の速さを加速して短時間で大域的極小値を探索しうる神経回路網モデルの大域的加速化学習法およびその装置を提供する。

【解決手段】 評価関数曲面上に局所的極小値が存在するような場合でも、大域的な極小値を探索することが可能となるとともに、局所的極小値を探索する際、谷などの条件数の大きい曲面が存在する場合、この領域において発生する重みの振動を抑制することにより収束を加速化し、大域的な極小値を探索する時間を短縮することができる。



【特許請求の範囲】

【請求項1】 神経回路網モデルの重みの初期値、学習率の初期値、学習率の増減率、平滑化微係数、慣性率、補正係数の初期値、極値探索および大域的探索の停止基準値、制約条件の係数、および大域的最適化に関する係数を与える第1のステップと、

学習すべき入力データの全パターンを神経回路網モデルに提示し、各入力パターンに対して神経回路網モデルの出力値を計算し、この出力値と教師データから計算される全パターンに対する評価基準値に基づいて現在の反復点（ k 反復目とする）における重みに関する評価関数の一次偏微分である勾配を計算する第2のステップと、

1から平滑化微係数を引いた値を前記第2のステップで求めた k 反復目の勾配に乘じ、この値と1反復前（ $k-1$ 反復目とする）の平滑化微分に平滑化微係数を乗じた値を加算し、 k 反復目における平滑化微分値を計算する第3のステップと、

前記第3のステップで求めた k 反復目の勾配および $k-1$ 反復目の平滑化微分値を各要素毎に個別に乘じた演算結果が正值の場合、学習率に学習率の増加率を加えて学習率を増加させ、演算結果が負値の場合、学習率に学習率の減少率を乘じて学習率を減少させ、各学習率を個別に更新する第4のステップと、

前記第2および第4のステップで求めた勾配および各重みに対応した個別の学習率を乘じた修正量を各要素毎に個別に計算し、この修正量に $k-1$ 反復目の重みの修正量に慣性率を乘じた修正量を各重みに個別に加える第5のステップと、

前記第2のステップで求めた k 反復目の勾配と $k-1$ 反復目の勾配から勾配の差分（ k 反復目における勾配の差分とする）を計算する第6のステップと、

前記第6のステップで求めた k 反復目の勾配の差分と1反復前の $k-1$ 反復目の勾配の差分の内積を計算し、この結果が負値の場合には、前記第5のステップで求めた重みの修正量に、更に k 反復目の勾配の差分と k 反復目の重みの修正量の内積値を k 反復目の勾配の差分の大きさ（ノルム）の2乗および2で除算した結果（スカラー量）に k 反復目の勾配の差分を乘じた補正量を重みに加えるかまたは前記第4のステップで求めた勾配と平滑化微分の乗算結果が正值の場合には補正量を付加しない第7のステップと、

k 反復目の勾配と $k-1$ 反復目の平滑化微分を乗算した結果が正值の場合は補正係数を減少し、乗算結果が負値の場合は補正係数を増加する第8のステップと、

前記第5および第7のステップで求めた重みの修正量を用いて重みを要素毎に個別に更新する第9のステップと、

前記第1ないし第9のステップの極値探索を行い、極値探索の停止基準を満たす場合には、極値探索を終了し、第11のステップを実行し、極値探索の停止基準を

満たさない場合には、前記第2のステップに戻り、極値探索の停止基準を満たすまで前記第2ないし第10のステップを反復実行する第10のステップと、

極値探索により探索された点である局所解を初期値として、局所解における評価関数の極値を取り除いた新たな評価関数を構成し、この評価関数を最小化することにより大域的探索を行う第11のステップと、

大域的探索の停止基準を満たす場合は終了し、満たさない場合で、かつ局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できた場合は、制約条件に関するパラメータを調整した後、前記第2のステップ以降の極値探索を実行する第12のステップと、

局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できない場合は、大域的最適化に関する係数を調整した後、前記第11のステップに戻り、前記第11ないし第13のステップの大域的探索を反復実行し、指定された回数だけ反復実行しても、大域的探索の停止条件を満たさない場合は処理を終了する第13のステップとを有することを特徴とする神経回路網モデルの大域的加速化学習法。

【請求項2】 神経回路網モデルの重みの初期値、学習率の初期値、学習率の増減率、平滑化微係数、慣性率、補正係数の初期値、極値探索および大域的探索の停止基準値、制約条件の係数、および大域的最適化に関する係数を与える第1のステップと、

学習すべき入力データの1パターンを神経回路網モデルに提示し、これにより神経回路網モデルの出力値を計算し、この出力値と1パターンの教師データから計算される評価基準値に基づき、現在の反復点（ k 反復目とする）における重みに関する評価関数の一次偏微分である勾配を計算する第2のステップと、

1から平滑化微係数を引いた値を前記第2のステップで求めた k 反復目の勾配に乘じ、この値と1反復前（ $k-1$ 反復目とする）の平滑化微分に平滑化微係数を乗じた値を加算し、 k 反復における平滑化微分値を各要素毎に計算する第3のステップと、

前記第3のステップで求めた k 反復目の勾配および $k-1$ 反復目の平滑化微分値を各要素毎に個別に乘じた演算結果が正值の場合、学習率に学習率の増加率を加えて学習率を増加させ、演算結果が負値の場合、学習率に学習率の減少率を乘じて学習率を減少させ、各学習率を個別に更新する第4のステップと、

前記第2および第4のステップで求めた勾配および各重みに対応した個別の学習率を乘じた修正量を各要素毎に個別に計算し、この修正量に $k-1$ 反復目の慣性項に慣性率を乘じた修正量を各重みに個別に加える第5のステップと、

前記第2のステップで求めた k 反復目の勾配と $k-1$ 反復目の勾配から勾配の差分（ k 反復目における勾配の差分とする）を計算する第6のステップと、

前記第3のステップで求めた k 反復目の勾配と $k-1$ 反復目の平滑化微分を各要素毎に計算し、この結果が負値の場合には、前記第5のステップで求めた重みの修正量に、現在の反復点における勾配の差分に固定値の補正係数を乗じた補正量を各要素毎に個別に各重みに加えるかまたは前記第4のステップで求めた勾配と平滑化微分の内積値が正值の場合には補正量を付加しない第7のステップと、

k 反復目の勾配と $k-1$ 反復目の平滑化微分を乗算した結果が正值の場合は補正係数を減少し、乗算結果が負値の場合は補正係数を増加する第8のステップと、前記第5および第7のステップで求めた重みの修正量を用いて重みを要素毎に個別に更新する第9のステップと、

極値探索の停止基準を満たす場合には、極値探索を終了して、第11のステップを実行し、極値探索の停止基準を満たさない場合には、前記第2のステップに戻り、極値探索の停止基準を満たすまで前記第2ないし第10のステップを反復実行する第10のステップと、

極値探索により探索された点である局所解を初期値として、局所解における評価関数の極値を取り除いた新たな評価関数を構成し、この評価関数を最小化することにより大域的探索を行う第11のステップと、

大域的探索の停止基準を満たす場合は終了し、満たさない場合で、かつ局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できた場合は、制約条件に関するパラメータを調整した後、前記第2のステップ以降の極値探索を実行する第12のステップと、局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できない場合は、大域的最適化に関する係数を調整した後、前記第11のステップに戻り、前記第11ないし第13のステップの大域的探索を反復実行し、指定された回数だけ反復実行しても、大域的探索の停止条件を満たさない場合は処理を終了する第13のステップとを有することを特徴とする神経回路網モデルの大域的加速化学習法。

【請求項3】 前記第3のステップは、1から平滑化微係数を引いた値を $k-1$ 反復目の平滑化微分値に乘じ、この値と平滑化微係数を勾配に乘じた値を加算し、 k 反復目における平滑化微分値を各要素毎に個別に計算し、前記第4のステップは、前記第3のステップで求めた k 反復目の平滑化微分および $k-1$ 反復目の平滑化微分値を各要素毎に個別に乘じ、この演算結果が正值の場合、学習率に学習率の増加率を加えて学習率を増加させ、演算結果が負値の場合、学習率に学習率の減少率を乗じて学習率を減少させるように学習率を要素毎に個別に更新し、

前記第7のステップは、前記第3のステップで求めた k 反復目の平滑化微分と $k-1$ 反復目の平滑化微分を各要素毎に乗算し、この結果が負値の場合には、前記第3の

ステップで求めた重みの修正量に、 k 反復目における勾配の差分に可変の補正係数を乗じた補正量を求めるかまたは k 反復目および $k-1$ 反復目の平滑化微分の乗算結果が正值の場合には補正量を計算せず、

前記第8のステップは、 k 反復目の平滑化微分と $k-1$ 反復目の平滑化微分を乗算した結果が正值の場合は補正係数を減少し、乗算結果が負値の場合は補正係数を増加することを特徴とする請求項1記載の神経回路網モデルの大域的加速化学習法。

10 【請求項4】 神経回路網モデルの重みの初期値、学習率の初期値、学習率の増減率、平滑化微係数、慣性率、補正係数の初期値、極値探索および大域的探索の停止基準値、制約条件の係数、および大域的最適化に関する係数を与える第1の手段と、

学習すべき入力データの全パターンを神経回路網モデルに提示し、各入力パターンに対して神経回路網モデルの出力値を計算し、この出力値と教師データから計算される全パターンに対する評価基準値に基づいて現在の反復点(k 反復目とする)における重みに関する評価関数の一次偏微分である勾配を計算する第2の手段と、

1から平滑化微係数を引いた値を前記第2の手段で求めた k 反復目の勾配に乘じ、この値と1反復前($k-1$ 反復目とする)の平滑化微分に平滑化微係数を乗じた値を加算し、 k 反復目における平滑化微分値を計算する第3の手段と、

前記第3の手段で求めた k 反復目の勾配および $k-1$ 反復目の平滑化微分値を各要素毎に個別に乘じた演算結果が正值の場合、学習率に学習率の増加率を加えて学習率を増加させ、演算結果が負値の場合、学習率に学習率の減少率を乗じて学習率を減少させ、各学習率を個別に更新する第4の手段と、

前記第2および第4の手段で求めた勾配および各重みに対応した個別の学習率を乗じた修正量を各要素毎に個別に計算し、この修正量に $k-1$ 反復目の重みの修正量に慣性率を乗じた修正量を各重みに個別に加える第5の手段と、

前記第2の手段で求めた k 反復目の勾配と $k-1$ 反復目の勾配から勾配の差分(k 反復目における勾配の差分とする)を計算する第6の手段と、

40 前記第6の手段で求めた k 反復目の勾配の差分と1反復前の $k-1$ 反復目の勾配の差分の内積を計算し、この結果が負値の場合には、前記第5の手段で求めた重みの修正量に、更に k 反復目の勾配の差分と k 反復目の重みの修正量の内積値を k 反復目の勾配の差分の大きさ(ノルム)の2乗および2で除算した結果(スカラー量)に k 反復目の勾配の差分を乗じた補正量を重みに加えるかまたは前記第4の手段で求めた勾配と平滑化微分の乗算結果が正值の場合には補正量を付加しない第7の手段と、
50 k 反復目の勾配と $k-1$ 反復目の平滑化微分を乗算した結果が正值の場合は補正係数を減少し、乗算結果が負値

の場合は補正係数を増加する第8の手段と、
前記第5および第7の手段で求めた重みの修正量を用いて重みを要素毎に個別に更新する第9の手段と、
前記第1ないし第9の手段の極値探索を行い、極値探索の停止基準を満たす場合には、極値探索を終了して、第11の手段による処理を実行し、極値探索の停止基準を満たさない場合には、前記第2の手段による処理に戻り、極値探索の停止基準を満たすまで前記第2ないし第10の手段による処理を反復実行する第10の手段と、
極値探索により探索された点である局所解を初期値として、局所解における評価関数の極値を取り除いた新たな評価関数を構成し、この評価関数を最小化することにより大域的探索を行う第11の手段と、
大域的探索の停止基準を満たす場合は終了し、満たさない場合で、かつ局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できた場合は、制約条件に関するパラメータを調整した後、前記第2の手段による処理以降の極値探索を実行する第12の手段と、
局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できない場合は、大域的最適化に関する係数を調整した後、前記第11の手段による処理に戻り、前記第11ないし第13の手段による大域的探索を反復実行し、指定された回数だけ反復実行しても、大域的探索の停止条件を満たさない場合は処理を終了する第13の手段とを有することを特徴とする神経回路網モデルの大域的加速化学習装置。

【請求項5】 神経回路網モデルの重みの初期値、学習率の初期値、学習率の増減率、平滑化微係数、慣性率、補正係数の初期値、極値探索および大域的探索の停止基準値、制約条件の係数、および大域的最適化に関する係

数を与える第1の手段と、
学習すべき入力データの1パターンを神経回路網モデルに提示し、これにより神経回路網モデルの出力値を計算し、この出力値と1パターンの教師データから計算される評価基準値に基づき、現在の反復点（ k 反復目とする）における重みに関する評価関数の一次偏微分である勾配を計算する第2の手段と、

1から平滑化微係数を引いた値を前記第2の手段で求めた k 反復目の勾配に乘じ、この値と1反復前（ $k-1$ 反復目とする）の平滑化微分に平滑化微係数を乗じた値を加算し、 k 反復における平滑化微分値を各要素毎に計算する第3の手段と、

前記第3の手段で求めた k 反復目の勾配および $k-1$ 反復目の平滑化微分値を各要素毎に個別に乘じた演算結果が正值の場合、学習率に学習率の増加率を加えて学習率を増加させ、演算結果が負値の場合、学習率に学習率の減少率を乗じて学習率を減少させ、各学習率を個別に更新する第4の手段と、

前記第2および第4の手段で求めた勾配および各重みに対応した個別の学習率を乗じた修正量を各要素毎に個別

に計算し、この修正量に $k-1$ 反復目の慣性項に慣性率を乗じた修正量を各重みに個別に加える第5の手段と、
前記第2の手段で求めた k 反復目の勾配と $k-1$ 反復目の勾配から勾配の差分（ k 反復目における勾配の差分とする）を計算する第6の手段と、

前記第3の手段で求めた k 反復目の勾配と $k-1$ 反復目の平滑化微分を各要素毎に計算し、この結果が負値の場合には、前記第5の手段で求めた重みの修正量に、現在の反復点における勾配の差分に固定値の補正係数を乗じた補正量を各要素毎に個別に各重みに加えるかまたは前記第4の手段で求めた勾配と平滑化微分の内積値が正值の場合には補正量を付加しない第7の手段と、

k 反復目の勾配と $k-1$ 反復目の平滑化微分を乗算した結果が正值の場合は補正係数を減少し、乗算結果が負値の場合は補正係数を増加する第8の手段と、

前記第5および第7の手段で求めた重みの修正量を用いて重みを要素毎に個別に更新する第9の手段と、
極値探索の停止基準を満たす場合には、極値探索を終了して、第11の手段による処理を実行し、極値探索の停止基準を満たさない場合には、前記第2の手段による処理に戻り、極値探索の停止基準を満たすまで前記第2ないし第10の手段による処理を反復実行する第10の手段と、

極値探索により探索された点である局所解を初期値として、局所解における評価関数の極値を取り除いた新たな評価関数を構成し、この評価関数を最小化することにより大域的探索を行う第11の手段と、

大域的探索の停止基準を満たす場合は終了し、満たさない場合で、かつ局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できた場合は、制約条件に関するパラメータを調整した後、前記第2の手段による処理以降の極値探索を実行する第12の手段と、
局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できない場合は、大域的最適化に関する係数を調整した後、前記第11の手段による処理に戻り、前記第11ないし第13の手段による大域的探索を反復実行し、指定された回数だけ反復実行しても、大域的探索の停止条件を満たさない場合は処理を終了する第13の手段とを有することを特徴とする神経回路網モデルの大域的加速化学習装置。

【請求項6】 前記第3の手段は、1から平滑化微係数を引いた値を $k-1$ 反復目の平滑化微分値に乘じ、この値と平滑化微係数を勾配に乘じた値を加算し、 k 反復目における平滑化微分値を各要素毎に個別に計算する手段を有し、

前記第4の手段は、前記第3の手段で求めた k 反復目の平滑化微分および $k-1$ 反復目の平滑化微分値を各要素毎に個別に乘じ、この演算結果が正值の場合、学習率に学習率の増加率を加えて学習率を増加させ、演算結果が負値の場合、学習率に学習率の減少率を乗じて学習率を

減少させるように学習率を要素毎に個別に更新する手段を有し、

前記第7の手段は、前記第3の手段で求めたk反復目の平滑化微分とk-1反復目の平滑化微分を各要素毎に乗算し、この結果が負値の場合には、前記第3の手段で求めた重みの修正量に、k反復目における勾配の差分に可変の補正係数を乗じた補正量を求めるかまたはk反復目およびk-1反復目の平滑化微分の乗算結果が正值の場合には補正量を計算しない手段を有し、

前記第8の手段は、k反復目の平滑化微分とk-1反復目の平滑化微分を乗算した結果が正值の場合は補正係数を減少し、乗算結果が負値の場合は補正係数を増加する手段を有することを特徴とする請求項4記載の神経回路網モデルの大域的加速化学習装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、ニューラルネットを用いた処理、非線形最適化法を用いた処理などに適用できる神経回路網モデルの大域的加速化学習法およびその装置に関し、更に詳しくは、学習データが提示される毎に重み振動を抑制した逐次学習を用いて学習時間の短縮化を図ると共に最適解を探索する際に乱数探索に変更して局所的極小値を回避し、大域的極小値に探索しうる神経回路網モデルの大域的加速化学習法およびその装置*

$$y = \text{sig}(x) = \frac{1}{1 - \exp(-\sum_{i=1}^L w_i x_i + b_i)} \quad \dots (1)$$

【外1】

今、神経回路モデル上にN個の重み $w = (w_1, w_2, \dots, w_N)$ があり、あ

る評価基準を最小化するような重みの値を推定するための学習則を考える。ここでは、評価基準の典型例として神経回路網モデルの出力値 $(O_j \ (j = 1, 2, \dots, M))$ 、M：出力層のユニットの番号)と学習用出力値で*

$$f(w) = \sum_{j=1}^M (T_j - O_j)^2 \quad \dots (2)$$

D.E.Rumelhartらは神経回路網モデルの学習法として重みの更新則に慣性項を付加した学習法を提案した。しかし、評価関数曲面が谷の形状を成す場合、学習過程において、重みが谷を横切る方向に振動するため、重みは、最適解が存在する谷を降下する方向に更新されず、収束速度が低下するという問題がある。この問題を解決するために、落合らは、重みの振動を抑制するKick Out法を

*に関する。

【0002】

【従来の技術】初めに、本発明が適用できる神経回路網モデルの一例を挙げ、これに対する学習法を説明する。ここでは、典型例として階層型神経回路網モデルを用いるが、再帰結合をもつ神経回路網モデルなど、他の構造をもつモデルにも適用できる。

【0003】階層型神経回路網モデルは、図2に示すように、1層の入力層、複数層の中間層、1層の出力層からなる層状のネットワークモデルであり、各層はユニット、重み、バイアスから構成される。ユニットは、前層のユニットの出力値 $(x_i \ (i = 1, 2, \dots, L, L: \text{前層のユニット数}))$ と重み $(w_i, i: \text{重みの番号})$ の積の総和、およびバイアス $(b_i, i: \text{ユニットの番号})$ を加算した値を入力値として受け、入力値にある非線形変換 $(f(\cdot))$ を施した値 (y) を出力し、この出力値を次層のユニットへ伝達する構造をもつ(図2、式(1))。但し、ここでは、入力層のユニットの入出力変換関数は、線形、入力層以外の層のユニットの非線形変換関数 $f(\cdot)$ は、典型例であるシグモイド関数を用いる(式(1))が、モデルに応じて他の変換関数を用いることも考えられる。

【0004】

【数1】

30※ある教師データ $(T_j \ (j = 1, 2, \dots, M))$ の残差二乗和(式(2))を用いる。

【0005】

【数2】

提案した(文献：落合慶広、戸田尚宏、臼井支朗：“重みの振動を抑制する階層型ニューラルネットワークの加速化—Kick Out法—”、電気学会論文誌、Vol.113-C, No.12, pp.1154-1162(1993)参照)。

【0006】この学習法を以下に示す。

【0007】

【数3】

但し、 k 反復目の重み、勾配（評価関数の一次偏微分）、学習率、平滑化係数、学習率の増加率、減少率を

$$w_k, g_k, \eta_k = (\eta_0, \eta_1, \dots, \eta_N), \theta, \kappa, \phi$$

とし、過去の反復点での勾配を考慮した平滑化微分を \bar{g} 、慣性率を α とする。

Kick Out法

重みの更新則：

$$\begin{aligned} w_{k+1} &= w_k + \Delta w_k \\ \Delta w_k &= -\text{diag}(\eta_k) g_k + \alpha \Delta w_{k-1} \end{aligned} \quad \dots (3)$$

$$\left\{ \begin{aligned} \Delta w_k &= -\text{diag}(\eta_k) g_k + \alpha \Delta w_{k-1} \\ &\quad \text{if } y_k^T y_{k-1} > 0 \\ \Delta w_k &= -\text{diag}(\eta_k) g_k + \alpha \Delta w_{k-1} - \frac{y_k^T \Delta w_k}{2 \|y_k\|^2} y_k \\ &\quad \text{if } y_k^T y_{k-1} < 0 \end{aligned} \right. \quad \dots (4)$$

【数4】

学習率の更新則：

$$\left\{ \begin{aligned} \eta_{k,i} &= \eta_{k-1,i} + \kappa & \text{if } \bar{\delta}_{k-1,i} \cdot g_{k,i} > 0 \\ \eta_{k,i} &= \eta_{k-1,i} \cdot \phi & \text{if } \bar{\delta}_{k-1,i} \cdot g_{k,i} < 0 \\ \eta_{k,i} &= \eta_{k-1,i} & \text{otherwise} \end{aligned} \right. \quad \dots (5)$$

$$\text{但し、} \bar{\delta}_k = (1-\theta) \cdot g_k + \theta \cdot \bar{\delta}_{k-1} \quad \dots (6)$$

しかしながら、Kick Out法は一括学習型の極値探索法であるため、評価関数曲面上に、局所的な極小値が存在する場合、局所的極小値に捕らわれる可能性が大きく、大域的最適解を探索することが困難である。

【0008】一方、評価関数曲面上に局所的極小値が存在する場合、探索過程において局所的極小値に収束するため、大域的極小値を探索できないという問題がある。

【0009】これを解決する学習則として、極値探索において局所的極小値が求められた場合、その局所解を中*

* 心としたペナルティ関数を構成し、極小値を消滅させた評価関数（（9）式）を構成し、これに基づいて新たな極小値を探索するダイナミックトンネルアルゴリズム（文献：Y.Yao: "Dynamic tunneling algorithm for global optimization", IEEE Trans. on Sys. Man and Cybern. SMC-19-5, pp.1222-1230(1989)参照）が提案されている。この方法は、2つのアルゴリズムから構成される。次のような最適化問題が与えられた場合、

【外2】

$$\begin{aligned} \min \quad & f(w) \\ \text{sub. } & g_j(w) \leq 0 \quad j=1, 2, \dots, m \end{aligned}$$

但し、 $f(w)$ は最小化すべき評価関数、 $g_j(w)$ は制約条件を表す。

トンネリングアルゴリズムでは、初めに、次式のような微分方程式を考え、これをルンゲクッタ法などの数値積分法を用いて解くことにより、局所解 w^* を求める。

(7)

特開平10-334070

11

12

【0010】

$$\frac{dw_i}{dt} = \frac{\partial f}{\partial w_i} - \sum_{j=1}^n k_j f^*(g_j(w)) \frac{\partial g_j(w)}{\partial w_i} \quad (i=1, 2, \dots, n) \quad \text{**【数5】}$$

... (7)

但し、

$$f^*(a) = \begin{cases} a & \text{if } a > 0 \\ 0 & \text{if } a < 0 \end{cases} \quad \text{... (8)}$$

【0011】

※10※【外3】

上記の極値探索により局所解 w^* が求められると、次式(9式)が成立するような解 w_{opti} を見つける最適化問題を考える。但し、 m は、見いだされた局所的極小値の数を、 n は w_{opti} が $H(w)$ の極になるような十分大きな数とする。

【0012】

★ ★【数6】

$$H(w) = \frac{f(w) - f(w_{optn})}{A \cdot B \cdots M} = 0 \quad \text{... (9)}$$

$$\text{但し、 } A = [(w - w_{opt1})^T (w - w_{opt1})]^{\eta_1}$$

$$B = [(w - w_{opt2})^T (w - w_{opt2})]^{\eta_2}$$

$$\vdots$$

$$M = [(w - w_{optn})^T (w - w_{optn})]^{\eta_n}$$

【0013】

☆ ☆【外4】

この方法では、極値探索により局所解 w^* が求められると、上式(10式)の微分方程式を構成し、これをルンゲクッタ法などを用いて解くことにより、更に評価関数値が小さい解を探索する。但し、 $f(w)$ は最小化すべき評価関数を表し、

【数7】

◆【外5】

$$\bar{f}(w) = f(w) - f(w^*)$$

とし、 w^* は極値探索により求められた局所解、 k_i はペナルティー係数、 $g_i(w)$ は制約条件を表す。

【0014】

* *【数8】

$$\frac{dw_i}{dt} = - \frac{\frac{\partial f}{\partial w_i}}{[(w - w^*)^T (w - w^*)]} - \sum_{j=1}^n k_j f^*(g_j(w)) \frac{\partial g_j(w)}{\partial w} \\ k f^*(\bar{f}(w)) \frac{\partial f}{\partial w_i} \quad (i=1, 2, \dots, n) \quad \text{... (10)}$$

前述したように、ダイナミックトンネルアルゴリズムでは、極値探索と大域探索を繰り返して用いることにより、大域的最適解を求めている。

【0015】しかしながら、この大域的最適化法で用いられている極値探索アルゴリズムは、原始的な最急降下法を用いており、収束が非常に遅く、また、数値積分に

よる常微分方程式の解を求めるのも収束速度の点では遅いという問題がある。

【0016】また、実際的には、大域的探索を行う際に調整しなければならないパラメータとして k_j 、 k があるが、この調整は経験的にしか調整することができず、その調整も非常に困難である。

【0017】

【発明が解決しようとする課題】Kick Out法は、評価関数曲面上における谷の形状に対して加速化の効果を発揮する極値探索型のアルゴリズムであり、Kick Out法のみでは、評価関数曲面上に局所的な極小値が存在する場合、これを回避することは困難である。

【0018】また、ダイナミックトネリングアルゴリズムは、大域的な最小値を探索するアルゴリズムであるが、収束が非常に遅いという問題がある。

【0019】更に、Kick Out法は、学習すべきパターンを全て提示した後に1回だけ重みを更新する一括学習法を用いており、重みの振動を補正するための補正項の係数は、全学習パターンを提示した後にしか求めることができない。従って、学習アルゴリズムを一反復するために要する時間は、パターン数の増加に対して指数関数的に増加し、学習時間も膨大になるという問題がある。これを改善するためには、学習パターンを1パターン提示する度に重みを更新する逐次学習を導入する必要があるが、一括学習型のKick Out法を単純に逐次学習型に変更しても、Kick Out法と同程度の加速効果は得られないという問題がある。

【0020】本発明は、上記に鑑みてなされたもので、その目的とするところは、評価関数曲面上に局所的極小値が存在しても大域的な極小値を探索することができるように学習の収束の速さを加速して短時間で大域的極小値を探索しうる神経回路網モデルの大域的加速化学習法およびその装置を提供することにある。

【0021】

【課題を解決するための手段】上記目的を達成するため、請求項1記載の本発明は、神経回路網モデルの重みの初期値、学習率の初期値、学習率の増減率、平滑化微係数、慣性率、補正係数の初期値、極値探索および大域的探索の停止基準値、制約条件の係数、および大域的最適化に関する係数を与える第1のステップと、学習すべき入力データの全パターンを神経回路網モデルに提示し、各入力パターンに対して神経回路網モデルの出力値を計算し、この出力値と教師データから計算される全パターンに対する評価基準値に基づいて現在の反復点(k 反復目とする)における重みに関する評価関数の一次偏微分である勾配を計算する第2のステップと、1から平滑化微係数を引いた値を前記第2のステップで求めた k 反復目の勾配に乘じ、この値と1反復前($k-1$ 反復目とする)の平滑化微分に平滑化微係数を乗じた値を加算し、 k 反復目における平滑化微分値を計算する第3のス

テップと、前記第3のステップで求めた k 反復目の勾配および $k-1$ 反復目の平滑化微分値を各要素毎に個別に乘じた演算結果が正值の場合、学習率に学習率の増加率を加えて学習率を増加させ、演算結果が負値の場合、学習率に学習率の減少率を乗じて学習率を減少させ、各学習率を個別に更新する第4のステップと、前記第2および第4のステップで求めた勾配および各重みに対応した個別の学習率を乗じた修正量を各要素毎に個別に計算し、この修正量に $k-1$ 反復目の重みの修正量に慣性率を乗じた修正量を各重みに個別に加える第5のステップと、前記第2のステップで求めた k 反復目の勾配と $k-1$ 反復目の勾配から勾配の差分(k 反復目における勾配の差分とする)を計算する第6のステップと、前記第6のステップで求めた k 反復目の勾配の差分と1反復前の $k-1$ 反復目の勾配の差分の内積を計算し、この結果が負値の場合には、前記第5のステップで求めた重みの修正量に、更に k 反復目の勾配の差分と k 反復目の重みの修正量の内積値を k 反復目の勾配の差分の大きさ(ノルム)の2乗および2で除算した結果(スカラー量)に k 反復目の勾配の差分を乗じた補正量を重みに加えるかまたは前記第4のステップで求めた勾配と平滑化微分の乗算結果が正值の場合には補正量を付加しない第7のステップと、 k 反復目の勾配と $k-1$ 反復目の平滑化微分を乗算した結果が正值の場合は補正係数を減少し、乗算結果が負値の場合は補正係数を増加する第8のステップと、前記第5および第7のステップで求めた重みの修正量を用いて重みを要素毎に個別に更新する第9のステップと、前記第1ないし第9のステップの極値探索を行い、極値探索の停止基準を満たす場合には、極値探索を終了して、第11のステップを実行し、極値探索の停止基準を満たさない場合には、前記第2のステップに戻り、極値探索の停止基準を満たすまで前記第2ないし第10のステップを反復実行する第10のステップと、極値探索により探索された点である局所解を初期値として、局所解における評価関数の極値を取り除いた新たな評価関数を構成し、この評価関数を最小化することにより大域的探索を行う第11のステップと、大域的探索の停止基準を満たす場合は終了し、満たさない場合で、かつ局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できた場合は、制約条件に関するパラメータを調整した後、前記第2のステップ以降の極値探索を実行する第12のステップと、局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できない場合は、大域的最適化に関する係数を調整した後、前記第11のステップに戻り、前記第11ないし第13のステップの大域的探索を反復実行し、指定された回数だけ反復実行しても、大域的探索の停止条件を満たさない場合は処理を終了する第13のステップとを有することを要旨とする。

【0022】請求項1記載の本発明にあっては、評価関

数曲面上に局所的極小値が存在するような場合でも、大域的な極小値を探索することが可能となるとともに、局所的極小値を探索する際、谷などの条件数の大きい曲面が存在する場合、この領域において発生する重みの振動を抑制することにより収束を加速化し、大域的な極小値を探索する時間を短縮することができる。

【0023】また、請求項2記載の本発明は、神経回路網モデルの重みの初期値、学習率の初期値、学習率の増減率、平滑化微係数、慣性率、補正係数の初期値、極値探索および大域的探索の停止基準値、制約条件の係数、および大域的最適化に関する係数を与える第1のステップと、学習すべき入力データの1パターンを神経回路網モデルに提示し、これにより神経回路網モデルの出力値を計算し、この出力値と1パターンの教師データから計算される評価基準値に基づき、現在の反復点（ k 反復目とする）における重みに関する評価関数の一次偏微分である勾配を計算する第2のステップと、1から平滑化微係数を引いた値を前記第2のステップで求めた k 反復目の勾配に乘以、この値と1反復前（ $k-1$ 反復目とする）の平滑化微分に平滑化微係数を乗じた値を加算し、 k 反復における平滑化微分値を各要素毎に計算する第3のステップと、前記第3のステップで求めた k 反復目の勾配および $k-1$ 反復目の平滑化微分値を各要素毎に個別に乘じた演算結果が正值の場合、学習率に学習率の増加率を加えて学習率を増加させ、演算結果が負値の場合、学習率に学習率の減少率を乗じて学習率を減少させ、各学習率を個別に更新する第4のステップと、前記第2および第4のステップで求めた勾配および各重みに対応した個別の学習率を乗じた修正量を各要素毎に個別に計算し、この修正量に $k-1$ 反復目の慣性項に慣性率を乗じた修正量を各重みに個別に加える第5のステップと、前記第2のステップで求めた k 反復目の勾配と $k-1$ 反復目の勾配から勾配の差分（ k 反復目における勾配の差分とする）を計算する第6のステップと、前記第3のステップで求めた k 反復目の勾配と $k-1$ 反復目の平滑化微分を各要素毎に計算し、この結果が負値の場合には、前記第5のステップで求めた重みの修正量に、現在の反復点における勾配の差分に固定値の補正係数を乗じた補正量を各要素毎に個別に各重みに加えるかまたは前記第4のステップで求めた勾配と平滑化微分の内積値が正值の場合には補正量を付加しない第7のステップと、 k 反復目の勾配と $k-1$ 反復目の平滑化微分を乗算した結果が正值の場合は補正係数を減少し、乗算結果が負値の場合は補正係数を増加する第8のステップと、前記第5および第7のステップで求めた重みの修正量を用いて重みを要素毎に個別に更新する第9のステップと、極値探索の停止基準を満たす場合には、極値探索を終了して、第11のステップを実行し、極値探索の停止基準を満たさない場合には、前記第2のステップに戻り、極値探索の停止基準を満たすまで前記第2ないし第10の

ステップを反復実行する第10のステップと、極値探索により探索された点である局所解を初期値として、局所解における評価関数の極値を取り除いた新たな評価関数を構成し、この評価関数を最小化することにより大域的探索を行う第11のステップと、大域的探索の停止基準を満たす場合は終了し、満たさない場合で、かつ局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できた場合は、制約条件に関するパラメータを調整した後、前記第2のステップ以降の極値探索を実行する第12のステップと、局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できない場合は、大域的最適化に関する係数を調整した後、前記第11のステップに戻り、前記第11ないし第13のステップの大域的探索を反復実行し、指定された回数だけ反復実行しても、大域的探索の停止条件を満たさない場合は処理を終了する第13のステップとを有することを要旨とする。

【0024】請求項2記載の本発明にあっては、評価関数曲面上に局所的極小値が存在するような場合でも、大域的な極小値を探索することが可能となるとともに、重みおよび学習率を個別に更新することが可能となり、多くの学習パターンを有する課題の学習または時系列データなどを用いた逐次学習などにおいて学習時間の大幅な短縮が可能となる。

【0025】更に、請求項3記載の本発明は、請求項1記載の発明において、前記第3のステップが、1から平滑化微係数を引いた値を $k-1$ 反復目の平滑化微分値に乘以、この値と平滑化微係数を勾配に乘じた値を加算し、 k 反復目における平滑化微分値を各要素毎に個別に計算し、前記第4のステップが、前記第3のステップで求めた k 反復目の平滑化微分および $k-1$ 反復目の平滑化微分値を各要素毎に個別に乘以、この演算結果が正值の場合、学習率に学習率の増加率を加えて学習率を増加させ、演算結果が負値の場合、学習率に学習率の減少率を乗じて学習率を減少させるように学習率を要素毎に個別に更新し、前記第7のステップが、前記第3のステップで求めた k 反復目の平滑化微分と $k-1$ 反復目の平滑化微分を各要素毎に乗算し、この結果が負値の場合には、前記第3のステップで求めた重みの修正量に、 k 反復目における勾配の差分に可変の補正係数を乗じた補正量を求めるかまたは k 反復目および $k-1$ 反復目の平滑化微分の乗算結果が正值の場合には補正量を計算せず、前記第8のステップが、 k 反復目の平滑化微分と $k-1$ 反復目の平滑化微分を乗算した結果が正值の場合は補正係数を減少し、乗算結果が負値の場合は補正係数を増加することを要旨とする。

【0026】請求項3記載の本発明にあっては、評価関数曲面上に局所的極小値が存在するような場合でも、大域的な極小値を探索することが可能となるとともに、学習時の収束の不安定性を減少させることができる。

【0027】請求項4記載の本発明は、神経回路網モデルの重みの初期値、学習率の初期値、学習率の増減率、平滑化微係数、慣性率、補正係数の初期値、極値探索および大域的探索の停止基準値、制約条件の係数、および大域的最適化に関する係数を与える第1の手段と、学習すべき入力データの全パターンを神経回路網モデルに提示し、各入力パターンに対して神経回路網モデルの出力値を計算し、この出力値と教師データから計算される全パターンに対する評価基準値に基づいて現在の反復点

(k反復目とする)における重みに関する評価関数の一次偏微分である勾配を計算する第2の手段と、1から平滑化微係数を引いた値を前記第2の手段で求めたk反復目の勾配に乘じ、この値と1反復前(k-1反復目とする)の平滑化微分に平滑化微係数を乗じた値を加算し、k反復目における平滑化微分値を計算する第3の手段と、前記第3の手段で求めたk反復目の勾配およびk-1反復目の平滑化微分値を各要素毎に個別に乘じた演算結果が正值の場合、学習率に学習率の増加率を加えて学習率を増加させ、演算結果が負値の場合、学習率に学習率の減少率を乗じて学習率を減少させ、各学習率を個別に更新する第4の手段と、前記第2および第4の手段で求めた勾配および各重みに対応した個別の学習率を乗じた修正量を各要素毎に個別に計算し、この修正量にk-1反復目の重みの修正量に慣性率を乗じた修正量を各重みに個別に加える第5の手段と、前記第2の手段で求めたk反復目の勾配とk-1反復目の勾配から勾配の差分(k反復目における勾配の差分とする)を計算する第6の手段と、前記第6の手段で求めたk反復目の勾配の差分と1反復前のk-1反復目の勾配の差分の内積を計算し、この結果が負値の場合には、前記第5の手段で求めた重みの修正量に、更にk反復目の勾配の差分とk反復目の重みの修正量の内積値をk反復目の勾配の差分の大きさ(ノルム)の2乗および2で除算した結果(スカラ一量)にk反復目の勾配の差分を乗じた補正量を重みに加えるかまたは前記第4の手段で求めた勾配と平滑化微分の乗算結果が正值の場合には補正量を付加しない第7の手段と、k反復目の勾配とk-1反復目の平滑化微分を乗算した結果が正值の場合は補正係数を減少し、乗算結果が負値の場合は補正係数を増加する第8の手段と、前記第5および第7の手段で求めた重みの修正量を用いて重みを要素毎に個別に更新する第9の手段と、前記第1ないし第9の手段の極値探索を行い、極値探索の停止基準を満たす場合には、極値探索を終了して、第11の手段による処理を実行し、極値探索の停止基準を満たさない場合には、前記第2の手段による処理に戻り、極値探索の停止基準を満たすまで前記第2ないし第10の手段による処理を反復実行する第10の手段と、極値探索により探索された点である局所解を初期値として、局所解における評価関数の極値を取り除いた新たな評価関数を構成し、この評価関数を最小化することにより大域的

探索を行う第11の手段と、大域的探索の停止基準を満たす場合は終了し、満たさない場合で、かつ局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できた場合は、制約条件に関するパラメータを調整した後、前記第2の手段による処理以降の極値探索を実行する第12の手段と、局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できない場合は、大域的最適化に関する係数を調整した後、前記第11の手段による処理に戻り、前記第11ないし第13の手段による大域的探索を反復実行し、指定された回数だけ反復実行しても、大域的探索の停止条件を満たさない場合は処理を終了する第13の手段とを有することを要旨とする。

【0028】請求項4記載の本発明にあつては、評価関数曲面上に局所的極小値が存在するような場合でも、大域的な極小値を探索することが可能となるとともに、局所的極小値を探索する際、谷などの条件数の大きい曲面が存在する場合、この領域において発生する重みの振動を抑制することにより収束を加速化し、大域的な極小値を探索する時間を短縮することができる。

【0029】また、請求項5記載の本発明は、神経回路網モデルの重みの初期値、学習率の初期値、学習率の増減率、平滑化微係数、慣性率、補正係数の初期値、極値探索および大域的探索の停止基準値、制約条件の係数、および大域的最適化に関する係数を与える第1の手段と、学習すべき入力データの1パターンを神経回路網モデルに提示し、これにより神経回路網モデルの出力値を計算し、この出力値と1パターンの教師データから計算される評価基準値に基づき、現在の反復点(k反復目とする)における重みに関する評価関数の一次偏微分である勾配を計算する第2の手段と、1から平滑化微係数を引いた値を前記第2の手段で求めたk反復目の勾配に乘じ、この値と1反復前(k-1反復目とする)の平滑化微分に平滑化微係数を乗じた値を加算し、k反復目における平滑化微分値を各要素毎に計算する第3の手段と、前記第3の手段で求めたk反復目の勾配およびk-1反復目の平滑化微分値を各要素毎に個別に乘じた演算結果が正值の場合、学習率に学習率の増加率を加えて学習率を増加させ、演算結果が負値の場合、学習率に学習率の減少率を乗じて学習率を減少させ、各学習率を個別に更新する第4の手段と、前記第2および第4の手段で求めた勾配および各重みに対応した個別の学習率を乗じた修正量を各要素毎に個別に計算し、この修正量にk-1反復目の慣性項に慣性率を乗じた修正量を各重みに個別に加える第5の手段と、前記第2の手段で求めたk反復目の勾配とk-1反復目の勾配から勾配の差分(k反復目における勾配の差分とする)を計算する第6の手段と、前記第3の手段で求めたk反復目の勾配とk-1反復目の平滑化微分を各要素毎に計算し、この結果が負値の場合には、前記第5の手段で求めた重みの修正量に、現在の

反復点における勾配の差分に固定値の補正係数を乗じた補正量を各要素毎に個別に各重みに加えるかまたは前記第4の手段で求めた勾配と平滑化微分の内積値が正值の場合には補正量を付加しない第7の手段と、 k 反復目の勾配と $k-1$ 反復目の平滑化微分を乗算した結果が正值の場合は補正係数を減少し、乗算結果が負値の場合は補正係数を増加する第8の手段と、前記第5および第7の手段で求めた重みの修正量を用いて重みを要素毎に個別に更新する第9の手段と、極値探索の停止基準を満たす場合には、極値探索を終了して、第11の手段による処理を実行し、極値探索の停止基準を満たさない場合には、前記第2の手段による処理に戻り、極値探索の停止基準を満たすまで前記第2ないし第10の手段による処理を反復実行する第10の手段と、極値探索により探索された点である局所解を初期値として、局所解における評価関数の極値を取り除いた新たな評価関数を構成し、この評価関数を最小化することにより大域的探索を行う第11の手段と、大域的探索の停止基準を満たす場合は終了し、満たさない場合で、かつ局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できた場合は、制約条件に関するパラメータを調整した後、前記第2の手段による処理以降の極値探索を実行する第12の手段と、局所解における評価関数値よりも更に小さな評価関数値を有する解が探索できない場合は、大域的最適化に関する係数を調整した後、前記第11の手段による処理に戻り、前記第11ないし第13の手段による大域的探索を反復実行し、指定された回数だけ反復実行しても、大域的探索の停止条件を満たさない場合は処理を終了する第13の手段とを有することを要旨とする。

【0030】請求項5記載の本発明にあっては、評価関数曲面上に局所的極小値が存在するような場合でも、大域的な極小値を探索することが可能となるとともに、重みおよび学習率を個別に更新することが可能となり、多くの学習パターンを有する課題の学習または時系列データなどを用いた逐次学習などにおいて学習時間の大幅な短縮が可能となる。

【0031】更に、請求項6記載の本発明は、請求項4記載の発明において、前記第3の手段が、1から平滑化微係数を引いた値を $k-1$ 反復目の平滑化微分値に乘以、この値と平滑化微係数を勾配に乘じた値を加算し、 k 反復目における平滑化微分値を各要素毎に個別に計算する手段を有し、前記第4の手段が、前記第3の手段で求めた k 反復目の平滑化微分および $k-1$ 反復目の平滑化微分値を各要素毎に個別に乘以、この演算結果が正值の場合、学習率に学習率の増加率を加えて学習率を増加させ、演算結果が負値の場合、学習率に学習率の減少率を乗じて学習率を減少させるように学習率を要素毎に個

別に更新する手段を有し、前記第7の手段が、前記第3の手段で求めた k 反復目の平滑化微分と $k-1$ 反復目の平滑化微分を各要素毎に乗算し、この結果が負値の場合には、前記第3の手段で求めた重みの修正量に、 k 反復目における勾配の差分に可変の補正係数を乗じた補正量を求めるかまたは k 反復目および $k-1$ 反復目の平滑化微分の乗算結果が正值の場合には補正量を計算しない手段を有し、前記第8の手段が、 k 反復目の平滑化微分と $k-1$ 反復目の平滑化微分を乗算した結果が正值の場合は補正係数を減少し、乗算結果が負値の場合は補正係数を増加する手段を有することを要旨とする。

【0032】請求項6記載の本発明にあっては、評価関数曲面上に局所的極小値が存在するような場合でも、大域的な極小値を探索することが可能となるとともに、学習時の収束の不安定性を減少させることができる。

【0033】

【発明の実施の形態】以下、図面を用いて本発明の実施の形態について説明する。

【0034】図1は、本発明の第1ないし第3の実施形態に適用される神経回路網モデルの大域的加速化学習法を実施する学習装置の構成を示すブロック図である。図1に示す学習装置は、入力部100、処理部200および出力部300から構成され、入力部100は、学習に用いる学習率などの各種変数などを読み込むためのファイル読み込み装置101、時系列データを計測する各種センサ102、TVカメラ103などからなる。処理部200は、変数初期化部201、データ提示部202、学習率更新部203、重み更新部204、勾配差分計算部205、補正項付加判断部206、補正量計算部207、重み補正部208、極値探索の収束判定部209、大域的探索部210、大域的探索収束判定部211、制約条件係数更新部212、大域的最適化係数更新部213からなる。

【0035】本発明の第1の実施形態に係る神経回路網モデルの大域的加速化学習法は、Kick Out法とトンネルアルゴリズムを融合させた学習法とし、トンネルアルゴリズムにおいて収束が遅いおよびパラメータ k の調整が困難であるという問題点を解決するために、次の手順により探索する。始めに、局所的極小値を見出すために(3)～(6)式のKick Out法を用いて極値探索を行う。そして、極値探索において局所的極小値が見出された後、この局所的極小値よりも更に小さな評価関数値をもつ点を見つけるために、次の手順に基づいた大域的探索を行う。

【0036】大域的探索を行う際的评价関数を次式のように定義する。

【0037】

【数9】

$$E(w) = \frac{f(w) - f(w_{opt_n})}{A \cdot B \cdots M} + \sum_{j=1}^n k_j \int_0^{g_j(w)} f^*(a) da + \lambda \hat{f}(w) \quad \cdots (11)$$

$$\text{但し、} A = [(w - w_{opt1})^T (w - w_{opt1})]^{q1}$$

$$B = [(w - w_{opt2})^T (w - w_{opt2})]^{q2}$$

$$\vdots$$

$$M = [(w - w_{optn})^T (w - w_{optn})]^{qn}$$

これに対して、大域的探索においてもKick Out法を適用する。 * 【0038】

$$\begin{cases} \Delta w_k = -\text{diag}(q_k) e_k + \alpha \Delta w_{k-1} & \text{if } z_k^T z_{k-1} > 0 \\ \Delta w_k = -\text{diag}(q_k) e_k + \alpha \Delta w_{k-1} - \frac{z_k^T \Delta w_k}{2 \|z_k\|^2} z_k & \text{if } z_k^T z_{k-1} < 0 \end{cases} \quad \cdots (12)$$

ここで、 e_k 、 z_k は次式 (13)、(14) で表される。

$$e_k = \frac{\partial E(w_k)}{\partial w_k} = \left(\frac{\partial E(w_k)}{\partial w_{k1}}, \frac{\partial E(w_k)}{\partial w_{k2}}, \dots, \frac{\partial E(w_k)}{\partial w_{kn}} \right) \quad \cdots (13)$$

$$z_k = e_k - e_{k-1} \quad \cdots (14)$$

また、ダイナミックトンネリングアルゴリズムでは、経路設定する。

験的にしか調整することができなかったパラメータ 【0040】

k_j 、 λ を次のように調整する。 【外6】

【0039】 (1) k_j 、 μ_0 、 $\alpha > 0$ 、 $0 < \beta < 1$ を※

(2) (3) ~ (6) 式のKick Out法を用いて $f(w)$ に対して最小化を行い

極値探索を実行する。これにより見いだされた極値を $w_{local.opt}$ とする。

(3) 次に、 $E(w)$ に対して最小化を行い、大域的探索を実行し、下記 (115) 式を満たせば終了する。

【数11】

$$g_j(w_{global.opt}) = \hat{f}(w_{global.opt}) = 0 \quad \cdots (15)$$

条件 (15) 式を満たさないで、大域的探索により、極値解における評価関

数値 $E(w_{local.opt})$ よりも更に小さな評価関数値をもつ点が見いだされた場

合、すなわち、
【数12】

$$* E(w_{\text{global.start}}) < E(w_{\text{local.opt}}) \quad \dots (16)$$

※ 【外8】
を満たす点を見出した場合、この点 $w_{\text{global.start}}$ を極値探索を開始する際の

開始点として、(6) 以下を実行する。(16) 式を満たす点を見出せなかった場合、(4) 以下を実行する。
※ 【0041】
【数13】

(4) 次式により係数 λ を更新する。

$$\lambda_{k+1} = \lambda_k + \mu_k \hat{f}(w_k) \quad \dots (17)$$

$k=0$ ならば、 $\mu_{\lambda,1} = \mu_{\lambda,0}$ とする。

(5) $|\hat{f}(w_k)| > \beta |\hat{f}(w_{k-1})|$ ならば、 $\mu_{\lambda,k+1} = \alpha \mu_{\lambda,k}$ 、それ以外ならば、

$\mu_{\lambda,k+1} = \mu_{\lambda,k}$ とおく。 $k=k+1$ として(3)へ戻る。

(6) 次式により係数 k_j を更新する。

$$k_{j,k+1} = k_{j,k} + \mu_k g_j(w_k) \quad \dots (18)$$

$k=0$ ならば、 $\mu_1 = \mu_0$ とする。

【0042】

★ ★ 【外9】
(7) $|g(w_k)| > \beta |g(w_{k-1})|$ ならば、 $\mu_{k+1} = \alpha \mu_k$ 、それ以外

ならば、 $\mu_{k+1} = \mu_k$ 、 $k=k+1$ として(2)へ戻る。しかし、大域的探索を指定された回数だけ反復実行しても、大域的探索の停止基準を満たさない場合は終了する。

【0043】本発明の第2の実施形態に係る大域的加速化学習法では、重みが振動する際、各要素ごとに個別に重みの値を補正する補正法を備えた逐次型の学習法であり、本学習法における補正項の計算方法は、重みの振動を判定するために、評価関数曲面上の大域的な勾配を表す平滑化微分、および勾配を利用して求めている。

【0044】記憶容量、計算量を極力低減させるため

30

に、学習率の更新基準、または重みの補正基準として k 反復目の勾配と $k-1$ 反復目の平滑化微分の積を用いている。更に、補正係数として計算値ではなく設定値を用いることにより、第1の実施形態で用いていた内積演算を排除し、重みの振動を各要素ごとに個別に補正することを可能にする。

【0045】第2の実施形態において、第1の実施形態の中で用いられているKick Out法に相当する部分の学習則を以下に示す。

【0046】

【数14】

第2の実施形態におけるKick Out法に相当する部分の学習則：

重みの更新則：

$$\begin{aligned} w_{k+1,i} &= w_{k,i} + \Delta w_{k,i} \\ \Delta w_{k,i} &= -\eta_{k,i} g_{k,i} + \alpha \Delta w_{k-1,i} \end{aligned} \quad \dots (19)$$

$$\left\{ \begin{array}{l} \Delta w_{k,i} = -\eta_{k,i} g_{k,i} + \alpha \Delta w_{k-1,i} \\ \qquad \qquad \qquad \text{if } g_{k,i} \cdot \overline{\delta_{k,i}} > 0 \\ \Delta w_{k,i} = -\eta_{k,i} g_{k,i} + \alpha \Delta w_{k-1,i} - \alpha_k y_{k,i} \\ \qquad \qquad \qquad \text{if } g_{k,i} \cdot \overline{\delta_{k,i}} < 0 \end{array} \right. \quad \dots (20)$$

学習率の更新則：

$$\left\{ \begin{array}{l} \eta_{k,i} = \eta_{k-1,i} + \kappa \quad \text{if } g_{k,i} \cdot \overline{\delta_{k-1,i}} > 0 \\ \eta_{k,i} = \eta_{k-1,i} \cdot \phi \quad \text{if } g_{k,i} \cdot \overline{\delta_{k-1,i}} < 0 \\ \eta_{k,i} = \eta_{k-1,i} \quad \text{otherwise} \end{array} \right. \quad \dots (21)$$

$$\text{但し、} \overline{\delta_{k,i}} = (1-\theta) \cdot g_{k,i} + \theta \cdot \overline{\delta_{k-1,i}} \quad \dots (22)$$

但し、上記は極値探索における学習則であり、大域的探索においては、下記の記号に置き換えた学習則を使用する。

【0047】

$$g_{k,i} \rightarrow e_{k,i} \quad \dots (23)$$

$$y_{k,i} \rightarrow z_{k,i} \quad \dots (24)$$

本発明の第3の実施形態の学習法では、学習を安定して収束するためには、学習率、および補正項付加の判断基準に用いる勾配などの評価関数曲面の情報として全パタ

ーンに対する大域的な情報を用いる必要があるため、
k、およびk-1反復目の平滑化微分の積を用いている。更に、補正係数を学習の状況に応じて変化させることにより、適切な補正を可能にする。

【0048】第3の実施形態で用いられる学習則を以下に示す。

【0049】

【数15】

第3の実施形態におけるKick Out法に相当する部分の学習則:

重みの更新則:

$$\begin{aligned} w_{k+1,i} &= w_{k,i} + \Delta w_{k,i} \\ \Delta w_{k,i} &= -\eta_{k,i} g_{k,i} + \alpha \Delta w_{k-1,i} \end{aligned} \quad \dots (25)$$

$$\left\{ \begin{aligned} \Delta w_{k,i} &= -\eta_{k,i} g_{k,i} + \alpha \Delta w_{k-1,i} \\ &\quad \text{if } \overline{\delta_{k-1,i}} \cdot \overline{\delta_{k,i}} > 0 \\ \Delta w_{k,i} &= -\eta_{k,i} g_{k,i} + \alpha \Delta w_{k-1,i} - a_k y_{k,i} \\ &\quad \text{if } \overline{\delta_{k-1,i}} \cdot \overline{\delta_{k,i}} < 0 \end{aligned} \right. \quad \dots (26)$$

学習率の更新則:

$$\left\{ \begin{aligned} \eta_{k,i} &= \eta_{k-1,i} + \kappa & \text{if } \overline{\delta_{k-1,i}} \cdot \overline{\delta_{k,i}} > 0 \\ \eta_{k,i} &= \eta_{k-1,i} \cdot \phi & \text{if } \overline{\delta_{k-1,i}} \cdot \overline{\delta_{k,i}} < 0 \\ \eta_{k,i} &= \eta_{k-1,i} & \text{otherwise} \end{aligned} \right. \quad \dots (27)$$

$$\text{但し、} \overline{\delta_{k,i}} = (1-\theta) \cdot g_{k,i} + \theta \cdot \overline{\delta_{k-1,i}} \quad \dots (28)$$

但し、上記は極値探索における学習則であり、大域的探索においては、下記の記号に置き換えた学習則を使用する。

【0050】

$$g_{k,i} \rightarrow e_{k,i} \quad \dots (29)$$

$$y_{k,i} \rightarrow z_{k,i} \quad \dots (30)$$

次に、図3および図4に示すフローチャートを参照して、作用を説明する。

【0051】図1において、入力部100からセンサ102、TVカメラ103などを用いて計測した時系列データなどを入力し、データ提示部202に学習用データを転送する(ステップS11)。従って、データ提示部202には時系列データが一時的に蓄えられる。また、ファイル読み込み装置101から、学習に必要な学習率などの値を読み込み、変数初期化部201に転送される(ステップS13)。

【0052】変数初期化部201は、神経回路網モデルの重み、および学習率などの初期値を設定し(ステップS15)、データ提示部202は、センサ102、TVカメラ103などから転送されてきた時系列データを1パターンずつ神経回路網モデルに提示し、前向き計算により、神経回路網モデルの出力値を計算し(ステップS17)、このモデル出力と教師データを用いて、予

め、与えられた評価基準値を計算する。この後、上記の評価基準値に基づいて神経回路網モデルの逆方向計算を行い、勾配を計算する(ステップS19)。

【0053】学習率変更部203は、勾配または平滑化微分の符号に基づいて学習率を更新し(ステップS21)、重み更新部204は、上述した学習率、勾配などを用いて重みを更新し、勾配などを勾配差分計算部205に転送する(ステップS23)。

【0054】勾配差分計算部205は、データ提示部202において求められた勾配を用いてk、k-1の反復点における勾配の差分を計算し、この情報などを補正項付加判断部206に転送する(ステップS25)。

【0055】補正項付加判断部206は、転送された勾配などに基づき、第1の実施形態においては、平滑化微分の積を、第2の実施形態においては、勾配と平滑化微分の積を計算し、この値が負値の場合、補正量計算部207に重みの値などを転送する。または、上記の値が正值の場合には、収束判定部209に重みの値などを転送する(ステップS27)。

【0056】補正量計算部207では、補正項付加判断部206の指令に従い、転送された重み、勾配の差分などから、重みの補正量を計算し、これを重み補正部208に転送する(ステップS29)。

30

40

50

【0057】重み補正部208は、転送された重みの補正量に基づき重みの値を補正し、補正した重みを収束判定部209に転送する（ステップS31）。

【0058】極値探索収束判定部209は、重み補正部208または重み更新部204から転送された重みを用いて、極値探索の停止基準が満たされているかを判定する（ステップS33）。

【0059】大域的探索部210は、新たに設定された評価関数 $E(w)$ に対して最小化を実行することにより、評価関数値 $E(w)$ がより小さくなる点を探索する（ステップS35）。

【0060】大域的探索収束判定部211は、大域的探索を実行した結果見出し出された点が大域的最適化の停止基準を満たしているかを判定する（ステップS37）。

【0061】制約条件係数更新部212は、大域的探索の結果、評価関数値がより小さな点が見出し出された場合には、制約条件の係数を更新する（ステップS39）。

【0062】大域的最適化係数更新部213は、大域的探索の結果、評価関数値がより小さな点が見出し出されなかった場合、大域的最適化の係数を更新する（ステップS41）。以上の処理を大域的探索を指定した回数だけ反復しても終了していない場合には、終了する。

【0063】上記各実施形態の学習装置には、（1）全ての学習率を同じ値に設定し、この値を学習過程において固定（学習率の増加率を0、学習率の減少率を1と設定）、または可変とした場合、（2）平滑化微係数を0と設定した場合、すなわち、重みの補正基準、学習率の更新基準に勾配のみを用いる場合、（3）慣性率の値を0に設定した場合も含まれる。

【0064】また、第2の実施形態の学習装置には、（1）補正係数を増加させる際、補正係数に正の値を加算する場合、または1より大きい値を乗算する場合、（2）補正係数を減少させる際、補正係数に、0より大きく、かつ、1未満の値を乗じる場合、または補正係数から正の値を減算する場合、（3）上記（1）、（2）を併用して、補正係数を増減する場合、（4）補正係数を増減せずに、一定値に設定する場合も含まれる。

【0065】

【発明の効果】以上説明したように、本発明によれば、評価関数曲面上に局所的極小値が存在するような問題に対して、大域的な極小値を探索することが可能となると

ともに、局所探索においては、評価関数曲面上で谷の構造が存在する場合には、収束を加速化し、短時間で大域的な極小値を探索することができる。

【0066】また、本発明によれば、評価関数曲面上に局所的極小値が存在するような場合でも大域的な極小値を探索することが可能となるとともに、少ない記憶領域および計算量により時系列データなどのように学習データが無数個存在する課題の学習を短時間で終了することができる。

【0067】更に、本発明によれば、評価関数曲面上に局所的極小値が存在するような場合でも大域的な極小値を探索することが可能となるとともに、時系列データなどのように学習データが無数個存在する課題の学習を短時間で終了し、かつ安定した収束が可能となる。

【図面の簡単な説明】

【図1】本発明の第1ないし第3の実施形態に適用される神経回路網モデルの大域的加速学習法を実施する学習装置の構成を示すブロック図である。

【図2】神経回路網モデルの一例を示す図である。

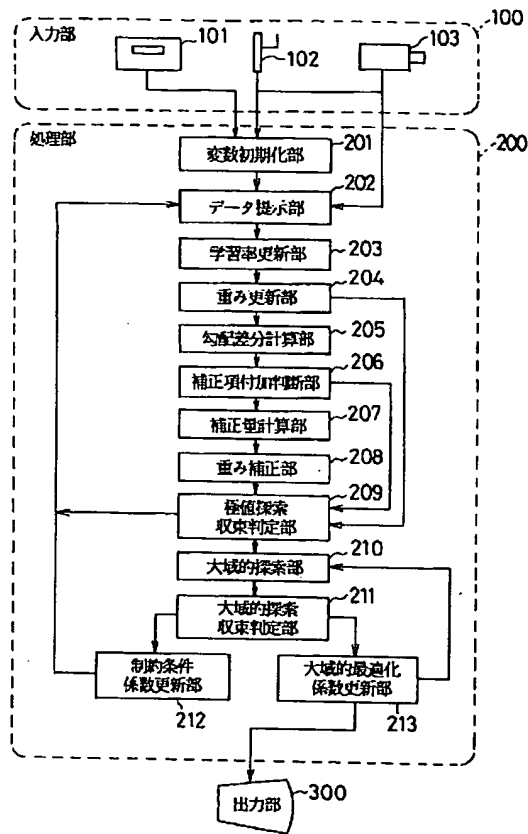
【図3】図1に示す神経回路網モデルの大域的加速学習法の手順の一部を示すフローチャートである。

【図4】図1に示す神経回路網モデルの大域的加速学習法の手順の残りの部分を示すフローチャートである。

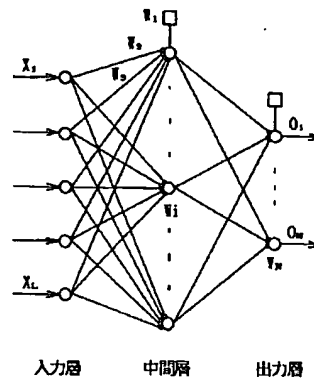
【符号の説明】

- 100 入力部
- 200 処理部
- 201 変数初期化部
- 202 データ提示部
- 203 学習率更新部
- 204 重み更新部
- 205 勾配差分計算部
- 206 補正項付加判断部
- 207 補正量計算部
- 208 重み補正部
- 209 極値探索収束判定部
- 210 大域的探索部
- 211 大域的探索収束判定部
- 212 制約条件係数更新部
- 213 大域的最適化係数更新部
- 300 出力部

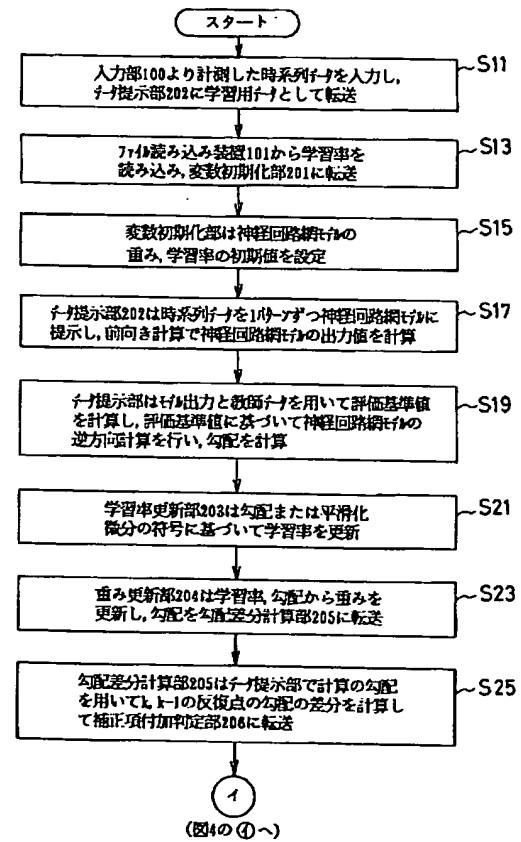
【図1】



【図2】



【図3】



【図4】

